The Limitations of X-ray Absorption Spectroscopy for Determining the Structure of Zinc Sites in Proteins. When Is a Tetrathiolate Not a Tetrathiolate?

Kimber Clark-Baldwin,[†] David L. Tierney,[†] Nandakumar Govindaswamy,[‡] Eric S. Gruff,[‡] Chongwoo Kim,[§] Jeremy Berg,[§] Stephen A. Koch,[‡] and James E. Penner-Hahn*,[†]

Contribution from the Departments of Chemistry, The University of Michigan, Ann Arbor, Michigan 48109-1055, State University of New York at Stonybrook, Stonybrook, New York 11974-3400, and The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received February 20, 1998

Abstract: X-ray absorption fine structure (XAFS) spectra have been measured for a series of structurally characterized zinc model complexes that mimic the zinc sites found in metalloproteins. These include both inorganic zinc coordination complexes and small zinc binding peptides. These data have been analyzed to determine the extent to which Zn XAFS can be used to determine reliably the ligation environment of the zinc. Because Zn−N and Zn−S XAFS oscillations are nearly out of phase over the accessible energy range, it is difficult to determine the relative number of sulfur and nitrogen scatterers, and in some cases, it is even difficult to determine *whether* a low-Z (N or O) ligand is bound in the presence of high-Z (S) ligands. We describe a protocol that, by controlling the number of variable parameters, can be used to obtain an accurate quantitation of the number of low-Z ligands. We also show that two of the variables that are often treated as freely adjustable parameters, the scale factor and shift in the threshold energy, can lead to erroneous results if not carefully controlled.

Zinc is essential for thousands of proteins in organisms spanning all known phyla.[1] It stabilizes the structure of a wide variety of biological molecules (proteins, nucleic acids, and cellular organelles) and plays a crucial role in diverse biological functions including gene replication and expression as well as protein biosynthesis and degradation. Zinc is the only metal known to be fundamental to the activity of at least one enzyme in each of the six classes of enzymes categorized by the International Union of Biochemistry.[2] A detailed understanding of these proteins, and the roles played by the zinc, requires structural characterization of the zinc sites.

Despite the undisputed need for information regarding local zinc-site structure, the majority of these environments remain uncharacterized. This is primarily due to the difficulty of studying Zn with conventional spectroscopic probes. Since biological zinc is present exclusively as the diamagnetic d^{10} Zn(II) ion, most of the common spectroscopic techniques are not useful. In many cases, the ligation of the zinc in proteins has been inferred on the basis of sequence similarity to one of the crystallographically characterized Zn proteins. Unfortunately, many zinc sites do not show significant sequence similarity to crystallographically characterized Zn proteins, and thus sequence alignment alone is often not sufficient to give the correct Zn ligation.[3,4] Even the observation of residue-dependent changes in Zn binding for site-directed mutants can give misleading results.[5,6]

The techniques of choice for characterizing biological zinc environments are X-ray crystallography and X-ray absorption fine structure spectroscopy (XAFS). X-ray crystallography has the advantage of providing information both on the zinc-site structure and on the overall tertiary structure of the protein, but is, naturally, limited to proteins which form diffraction-quality crystals. There are now nearly 400 Zn-containing structures deposited in the Protein Data Bank, although some of these entries represent multiple structures determinations on either the same or mutant proteins, and thus do not provide new Zn-site information. In recent years, new Zn-containing structures have been deposited at the rate of 80−90 per year. Nevertheless, only a small minority of the thousands of known Zn proteins have been crystallographically characterized.

The majority of the crystallographically characterized zinc proteins have four-coordinate Zn sites with pseudotetrahedral geometries. When only three amino acids are available to

* To whom correspondence should be addressed.
[†] The University of Michigan.
[‡] State University of New York at Stonybrook.
[§] The Johns Hopkins University School of Medicine.

(1) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647−5659.
(2) Vallee, B. L.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 220−224.
(3) Giedroc, D. P.; Qiu, H.; Khan, R.; King, G. C.; Chen, K. *Biochemistry* **1992**, *31*, 765−774.
(4) Severne, Y.; Wieland, S.; Schaffner, W.; Rusconi, S. *EMBO J.* **1988**, *7*, 2503−2508.
(5) Chong, S.; Curnow, A. W.; Huston, T. J.; Garcia, G. A. *Biochemistry* **1995**, *34*, 3694−3701.
(6) Romier, C.; Reuter, K.; Suck, D.; Ficner, R. *EMBO J.* **1996**, *15*, 2850−2857.
(7) Huheey, J. E. *Inorganic Chemistry*, 3rd ed.; Harper & Row: New York, 1983; p 314.
(8) Summers, M. F.; Henderson, L. E.; Chance, M. R.; Bess, J. W., Jr.; South, T. L.; Blake, P. R.; Sagi, I.; Perez-Alvarado, G.; Sowder, R. C., III; Hare, D. R.; Arthur, L. O. *Protein Sci.* **1992**, *1*, 563−574.
(9) Guo, J.; Wang, S.; Dong, J.; Qiu, H.; Scott, R. A.; Giedroc, D. P. *J. Am. Chem. Soc.* **1995**, *117*, 9437−40.
(10) Hubbard, S. R.; Bishop, W. R.; Kirschmeier, P.; George, S. J.; Cramer, S. P.; Hendrickson, W. A. *Science* **1991**, *254*, 1776−9.

S0002-7863(98)00580-0 CCC: $15.00
Published on Web 08/11/1998

coordinate to the zinc, water binds to complete the tetrahedral coordination sphere.[1] Since zinc is intermediate between hard and soft acids, it shows little preference for binding hard bases (e.g., oxygen) over soft bases (e.g., sulfur).[7] Consequently, zinc ligation spheres often contain a mixture of nitrogen, oxygen, and sulfur ligands in the form of histidine, aspartate and glutamate, and cysteine residues, respectively.

In contrast to crystallography, XAFS provides only limited radial structure information. Typically only scatterers within 4 Å of the Zn are detectable and little if any angular detail can be obtained. The decided advantage of XAFS, however, is the fact that it can be used to study noncrystalline samples. This has resulted in a plethora of articles describing the structural characterization of zinc sites in metalloproteins by XAFS. However, careful analysis of the uncertainties in Zn XAFS have led to the conclusion that there are often several possible solutions that are consistent with the XAFS data.[8] In particular, ZnS_4 and ZnS_3N sites often cannot be distinguished.[9,10] As a consequence, many studies have been limited to confirming that a previously deduced structure is consistent with the observed Zn XAFS data.

Worse still, the interference between M−S and M−N scattering, together with the lack of a widely accepted protocol for XAFS data analysis, has resulted in several examples of incorrect discrimination between S and N ligands. In GAL4, the tetrathiolate Zn site[11] was incorrectly modeled as having ZnS_3O ligation.[12] The opposite error, overestimation of the fraction of sulfur ligands, occurred for the Zn site in aminolevulinate dehydratase, where the $ZnCys_3(O)$ site[13] was modeled as having ZnS_4 ligation,[14] and in the Rieske [2Fe-2S] cluster, where the FeS_3N_1 average site[15] was modeled as having $FeS_{3.5}N_{0.5}$ ligation.[16]

To define the limitations of XAFS for the characterization of metal sites having mixed sulfur and nitrogen/oxygen ligation, we have undertaken an XAFS study of two sets of structurally characterized zinc model compounds. The models include S_4, S_3N, and S_2N_2 environments and thus mimic many of the zinc sites found in biology. We find that there *are* XAFS analysis protocols that can be used to distinguish reliably between these different Zn sites. We also find that it is surprisingly easy to obtain misleading results for the Zn site structure and to, as the title suggests, turn an authentic tetrathiolate into a site that does not appear, by XAFS, to have tetrathiolate ligation. Application of our data analysis protocol to obtain new insights into the structure and function of the Zn sites in cobalamin-dependent and cobalamin-independent methionine synthase is described in the accompanying paper.[17]

Experimental Section

Sample Preparation. The Zn model compounds $[Et_4N]_2[Zn(S\text{-}2\text{-}Ph\text{-}C_6H_4)_4]\cdot 2CH_3CN$,[18] $Pr_4N[Zn(S\text{-}2,3,5,6\text{-}Me_4C_6H)_3(1\text{-methylimidazole})]$,[19] and $[Zn(S\text{-}2,3,5,6\text{-}Me_4C_6H)_2(1\text{-methylimidazole})_2]$[20,21] were prepared as previously described. Samples were ground to a fine powder and diluted with BN. The diluted powder was packed into a thin (0.5 mm) Al sample cell with Kapton windows.

Table 1. Data Collection Parameters

sample	beamline[a]	temp (K)	harmonic rejection[b]
$[Zn(S\text{-}2,3,5,6\text{-}Me_4C_6H)_2(1\text{-Me-imid})_2]$	NSLS, X-9B	40	mirror
$[Zn(S\text{-}2,3,5,6\text{-}Me_4C_6H)_3(1\text{-Me-imid})]^{1-}$	NSLS, X-9B	40	mirror
$[Zn(S\text{-}2,3,5,6\text{-}Me_4C_6H)_4]^{2-}$	NSLS, X-9A	77	mirror
$Zn(cys)_2(his)_2$	SSRL, 7−3	10	detune
	NSLS, X-9B	40	mirror
$Zn(cys)_3(his)$	SSRL, 7−3	10	detune
	NSLS, X-9B	40	mirror
$Zn(cys)_4$	SSRL, 7−3	10	detune
	NSLS, X-9B	40	mirror

[a] NSLS is the National Synchrotron Light Source; SSRL is the Stanford Synchrotron Radiation Laboratory. [b] Harmonics were rejected either by using a focusing mirror or by detuning the monochromator crystals by 50%.

Metal binding peptides with the consensus sequence $ProTyrLysCys_4\text{-}ProGluCys_7GlyLysSerPheSerGlnLysSerAspLeuValLysXaa_{20}GlnArg\text{-}ThrYaa_{24}ThrGly$ were prepared and purified as previously described.[22,23] These peptides have coordination sites Cys_2His_2 (Xaa = Yaa = His), Cys_3His (Xaa = His, Yaa = Cys), and Cys_4 (Xaa = Yaa =Cys). Each peptide sample was treated with 0.9 equiv of zinc in 100 mM HEPES, 50 mM NaCl, pH 7.0 buffer. Solutions were loaded into Lucite sample cells with 6.25 μm polypropylene windows and frozen rapidly in liquid nitrogen. Two independent sets of peptide samples were prepared, having final Zn concentrations of ca. 1 and 8 mM, respectively.

X-ray Absorption Measurements. The experimental conditions are summarized in Table 1. All data were measured at 77 K or lower to decrease the dynamic disorder in the samples. The data for the inorganic zinc models were measured in transmission mode with an ionization detector. The zinc peptide model data were measured as Kα fluorescence excitation spectra with a 13-element solid-state Ge detector array. The total incident count rate of the Ge detector was kept below 50 kHz per channel to minimize deadtime distortion. The windowed (Zn Kα fluorescence) count rates were ca. 1.5−2.0 kHz per channel. All spectra were measured with a Si(220) double crystal monochromator. Other experimental parameters are summarized in Table 1.

EXAFS spectra were measured by using 10 eV steps with 1 s integration times in the preedge region (9300−9600 eV), ca. 0.5 eV steps in the edge region (9630−9700 eV), and 0.05 $Å^{-1}$ steps in the EXAFS region, using integration times increasing up to ca. 15 s at k = 13 (peptides) or 17 $Å^{-1}$ (inorganic models). Typically, two scans and six scans were collected and averaged for transmission and fluorescence data, respectively. The total integration time per scan was 35−40 min, with a total exposure time of approximately 1−2 h for the inorganic zinc models and 3−4 h for the zinc peptide models. Comparison of the first and last scan for each sample showed no evidence of radiation damage.

Data from individual fluorescence detector channels for each scan were examined to confirm the absence of artifacts before averaging. For the peptide data, each scan had 10−12 useful channels, giving a total of $2-3 \times 10^6$ Zn Kα fluorescence counts at $k = 13$ $Å^{-1}$. Energy calibration was accomplished by simultaneous collection of the absorption spectrum of a Zn metal foil, with the first inflection point of the foil assigned as 9659 eV.

(11) Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C. *Nature* **1992**, *356*, 408−414.
(12) Povey, J. F.; Diakun, G. P.; Garner, C. D.; Wilson, S. P.; Laue, E. D. *FEBS Lett.* **1990**, *266*, 142−146.
(13) Erskine, P. T.; Senior, N.; Awan, S.; Lambert, R.; Lewis, G.; Tickle, I. J.; Sarwar, M.; Spencer, P.; Thomas, P.; Warren, M. J.; Shoolingin-Jordan, P. M.; Wood, S. P.; Cooper, J. B. *Nature Struct. Biol.* **1997**, *4*, 1025−1031.
(14) Dent, A. J.; Beyersmann, D.; Block, C.; Hasnain, S. S. *Biochemistry* **1990**, *29*, 7822−7828.
(15) Iwata, S.; Saynovits, M.; Link, T. A.; Michel, H. *Structure* **1996**, *4*, 567−579.
(16) Powers, L.; Schägger, H.; von Jagow, G.; Smith, J.; Chance, B.; Ohnishi, T. *Biochim. Biophys. Acta* **1989**, *975*, 293−298.
(17) Peariso, K.; Goulding, C. W.; Huang, S.; Matthews, R. G.; Penner-Hahn, J. E. *J. Am. Chem. Soc.* **1998**, *120*, 8410−8416.
(18) Silver, A.; Koch, S. A.; Millar, M. *Inorg. Chim. Acta* **1993**, *205*, 9−14.
(19) Gruff, E. S.; Koch, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8762−3.
(20) Corwin, D. T., Jr; Koch, S. A. *Inorg. Chem.* **1988**, *27*, 493−6.
(21) Corwin, D. T., Jr; Gruff, E. S.; Koch, S. A. *J. Chem. Soc., Chem. Commun.* **1987**, 966−7.
(22) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4518−4523.
(23) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937−940.

The EXAFS data were extracted by first subtracting the background, using a first-order polynomial over the preedge region and a two-region cubic spline above the edge, followed by normalization to a Victoreen polynomial.[24] The data were converted to k space, where $k = [(8\pi^2 m_e/h^2)(E - E_0)]^{1/2}$, using a threshold energy of 9675 eV. Except where noted, the EXAFS data were Fourier transformed over the region $k = 2-12$ Å^{-1} followed by back-transformation over $R = 0.7-3.0$ Å. The resulting filtered EXAFS data were fit to eq 1 with a nonlinear least-squares algorithm.

$$\chi(k) = \sum_i \frac{N_i S_i(k) F_i(k)}{k R_i^2} \exp(2k^2\sigma_i^2) \exp\left(\frac{-2R_i}{\lambda}\right) \sin(2kR_i + \phi_i(k)) \quad (1)$$

In eq 1, N is the number of scatterers in shell i at a distance R from the absorber, σ is the root-mean-square variation in R, $F_i(k)$ is the backscattering amplitude of the absorber, λ is the photoelectron mean-free path, and $\phi(k)$ is the phase shift. $S_i(k)$ is a scale factor that is in principle unique to the absorber-scatterer pair. Equation 1 is summed over all absorber−scatterer pairs. Ab initio parameters, from the program FEFF version 5.01,[25,26] were used for the amplitude and phase functions, $F_i(k)\exp(-2R_i/\lambda)$ and $\phi_i(k)$. Calculations were done for a zinc−nitrogen interaction at 2.05 Å and a zinc−sulfur interaction at 2.35 Å. Parameters were calibrated by fitting EXAFS data for [Zn(*N*-methylimidazole)$_4$]$^{2+}$ for zinc−nitrogen and [Zn(SPh)$_4$]$^{2-}$ for zinc−sulfur, allowing R, σ, $S_i(k)$, and E_0 to float. The calibrated values are $S_i(k) = 0.85$ and $\Delta E_0 = 9$ eV for the zinc−nitrogen model and $S_i(k) = 1.02$ and $\Delta E_0 = 9$ eV for the zinc−sulfur model. Subsequent fits to the model data allowed only R and σ for each shell to vary while holding all other parameters fixed.

Results

Qualitative Differences between Spectra. The normalized X-ray absorption near edge structure (XANES) spectra for the inorganic models and for the peptide models are shown in Figures 1a and 1b, respectively. The energy resolution for all XANES spectra was identical. From Figure 1, it is clear the spectra show systematic changes as a function of nitrogen ligation. Within a set of models, the absorption edge energy (as measured by the first inflection point) increases and the edge becomes broader with increasing number of nitrogen ligands. The first peak in the XANES is most intense for the ZnS_4 models and weakest for the ZnS_2N_2 models. In principle, this sensitivity to ligation means that XANES spectra could be used to determine the Zn ligation. Unfortunately, comparison of the XANES spectra between the inorganic and the peptide model sets suggests that in practice this will be difficult. For example, the differences between the peptide and the inorganic ZnS_4 spectra are larger than the differences between the peptide ZnS_4 and ZnS_3N spectra. These differences may be due in part to differences in ligation type (e.g. arene- vs alkylthiolate). However, it seems more likely that they reflect minor variations in the Zn-site geometry. Comparable variations are seen in the published XANES spectra for nominally identical Zn sites.[9,27,28] To the extent that geometry-dependent variations are important, it will be difficult to use XANES spectra to determine ligation, although they are likely to be useful for detecting *changes* in a Zn site (e.g. on binding of substrate).

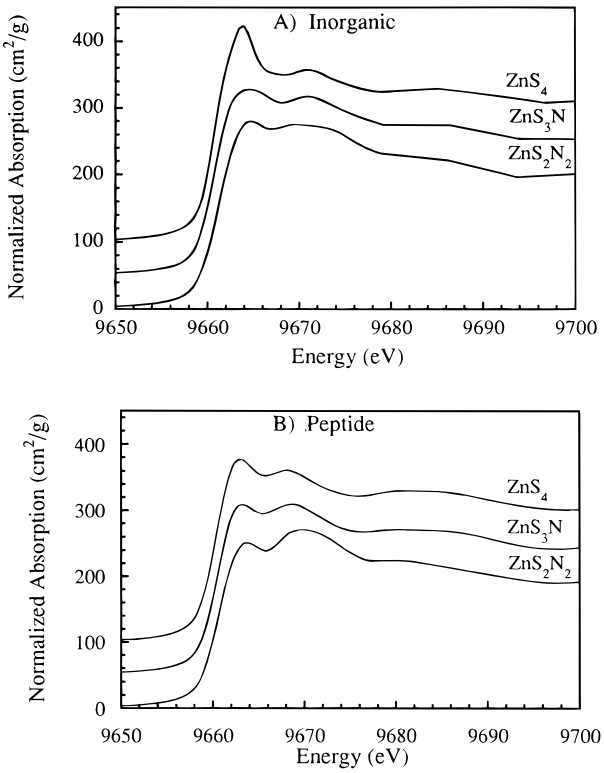

Figure 1. Normalized Zn XANES spectra: (A) inorganic model complexes and (B) zinc binding peptides. For both plots, spectra are, from the top, ZnS_4, ZnS_3N, and ZnS_2N_2. All spectra are on the same scale and offset vertically for clarity.

The extended X-ray absorption fine structure (EXAFS) spectra for the inorganic and peptide models and the corresponding Fourier transforms (FTs) are shown in Figure 2. For both sets of models, the amplitude of the EXAFS oscillations, especially at high k, is directly proportional to the number of sulfur ligands. It is apparent from Figure 2 (insets) that there are only small frequency differences between the spectra. This reflects the fact that Zn−S scattering dominates the EXAFS for all of the models. The peptide data were analyzed over the range $k = 2-12$ Å^{-1}, which is typical of Zn EXAFS for proteins. In this case, all of the FTs show a single peak whose magnitude (roughly 15, 23, and 30 for ZnS_2N_2, ZnS_3N, and ZnS_4 environments, respectively) is directly proportional to the number of sulfur ligands. The height of the main peak for the inorganic models is similar to the height for the analogous peptide models.

For the inorganic models, data were measured to higher k. Fourier transformation over $k = 2-16$ Å^{-1} allows partial resolution of the nitrogen and sulfur shells (Figure 2, top), although the ability to resolve the Zn−N and Zn−S shells depends dramatically on the upper limit of k as illustrated in Figure 3. Two peaks at approximately 1.6 and 1.9 Å are clearly visible in the FT of the ZnS_2N_2 data when $k_{max} = 16$ Å^{-1}, but are not resolved when $k_{max} \leq 13$ Å^{-1}. Similar effects, with somewhat worse resolution, are seen for the inorganic ZnS_3N model.

Two shells of *identical* scatterers should be resolvable in an EXAFS spectrum when their bond lengths differ by

$$\Delta R = \pi/2\Delta k \quad (2)$$

For the ZnS_2N_2 model, with $R_{Zn-N} = 2.05$ Å and $R_{Zn-S} = 2.30$ Å, eq 2 suggests that the shells should be resolvable for $k_{max} \geq 6.3$ Å^{-1}. Clearly, this is not the case. The difficulty is that eq

(24) *International Tables for X-ray Crystallography*; MacGillavry, C. H., Rieck, G. D., Ed.; D. Reidel: Boston, 1985; Vol. 3, pp 171−174.
(25) Rehr, D. L.; Zabinsky, A. *J. Am. Chem. Soc.* **1991**, *113*, 5135−5140.
(26) Rehr, J. J.; Zabinsky, S. I.; Albers, R. C. *Phys. Rev. Lett.* **1992**, *69*, 3397.
(27) George, G. N.; Pickering, I. J.; Prince, R. C.; Zhou, Z. H.; Adams, M. W. W. *J. Biol. Inorg. Chem.* **1996**, *1*, 226−230.
(28) Zeng, Q.; Lewis, L. M.; Colangelo, C. M.; Dong, J.; Scott, R. A. *J. Biol. Inorg. Chem.* **1996**, *1*, 162−168.

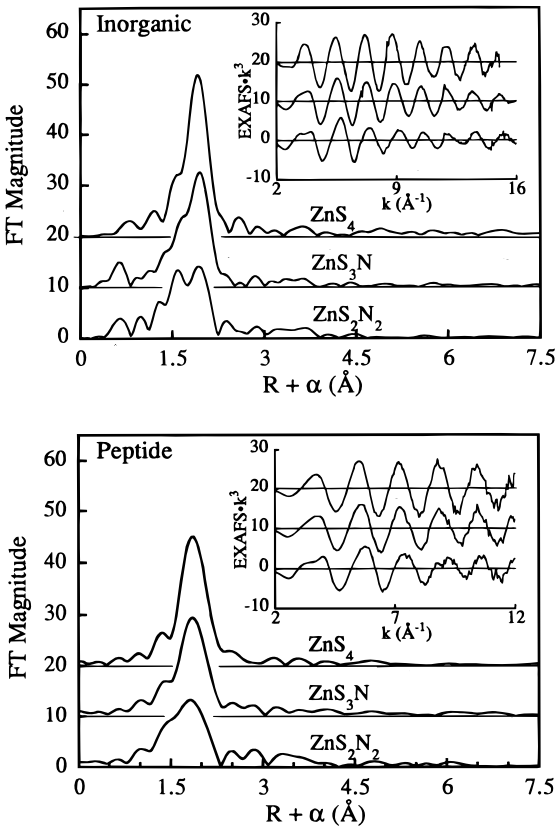

Figure 2. EXAFS spectra (insets) and Fourier transforms of the EXAFS spectra for the models. (Top) Inorganic models over the range $k = 2-16$ Å^{-1} (FT over $2-15$ Å^{-1} for ZnS_4). (Bottom) Zinc binding peptides over the range $k = 2-12$ Å^{-1}. The spectra appear in the same order as in Figure 1, offset vertically for clarity.

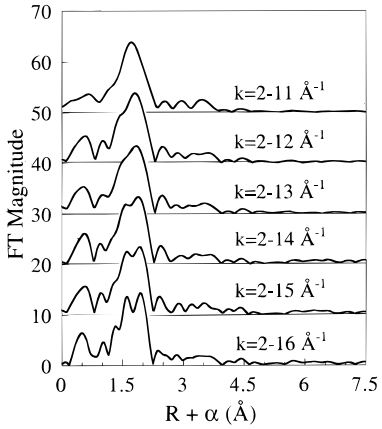

Figure 3. Fourier transform of the inorganic ZnS_2N_2 complex as a function of k_{max}. Data are plotted on the same scale, offset vertically.

2 assumes identical scatterers, while Zn−S and Zn−N EXAFS differ in both amplitude and phase. As illustrated by Figure 3, these differences complicate the problem of distinguishing different Zn ligation environments.

Quantitative Analyses. Quantitative curve fitting results are given in Table 2. The goodness of fit, $\mathcal{E}$, is defined as the root-mean-square error between the data and the fit (eq 3). In eq 3, $\chi_{obs}(k)$ represents the experimental data and $\chi_{calc}(k)$ is the calculated EXAFS.

$$\mathcal{E} = \sqrt{\sum_{i=1}^{N} \frac{k^6(\chi_{obs}(k_i) - \chi_{calc}(k_i))^2}{N}} \qquad (3)$$

Table 2. Curve Fitting Results[a]

sample	coordination no.	R_{ab} (Å)[b]	σ^2 [c]	$\mathcal{E}$ [d]	F [e]
ZnS_2N_2 inorganic	2N	2.04	4.3	0.3326	5.98
	2S	2.25	4.1		
	2S	2.23	2.7	0.6149	0.14
	2S	2.28	16.9		
ZnS_2N_2 peptide	2N	2.03	2.8	0.3481	4.75
	2S	2.26	2.9		
	2S	2.20	6.9	0.6085	0.02
	2S	2.29	5.7		
ZnS_3N inorganic	1N	2.12	9.6	0.359	2.84
	3S	2.28	4.0		
	2S	2.28	2.4	0.4674	0.74
	2S	2.30	9.2		
ZnS_3N peptide	1N	2.11	11.7	0.3408	2.68
	3S	2.28	3.9		
	2S	2.28	2.3	0.4054	1.23
	2S	2.30	14.2		
ZnS_4 inorganic	4S	2.30	3.8	0.5526	
	2S	2.26	0.0	0.4821	0.71
	2S	2.36	0.8		
ZnS_4 peptide	4S	2.30	3.9	0.5077	
	2S	2.26	2.4	0.3616	2.21
	2S	2.36	1.9		

[a] Fits to filtered data. Unfiltered data gave identical metric parameters. [b] Absorber−scatterer distance. [c] Debye−Waller factor $\times 10^3$. [d] Root-mean-square deviation (see text). [e] F values relative to single shell 4S fit, assuming $\Delta k = 10$ Å^{-1} and $\Delta R = 1.5$ Å ($F = [(\mathcal{E}_2^2 - \mathcal{E}_1^2)/2]/[\mathcal{E}_2^2/\nu]$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are for 1 and 2 shell fits, see text). To be significant at $P = 0.05$, F should be larger than ≈5.8.

Owing to the nature of least-squares fitting, $\mathcal{E}$ is almost always smaller for a two-shell refinement than for the corresponding single-shell fit simply due to the increase in the number of variable parameters. It is important to note, therefore, that an improvement in $\mathcal{E}$ does not necessarily indicate a better description of the data.

In principle, standard statistical methods should be useful to test the validity of including an additional term (i.e., additional shell of scatterers) in eq 1. An approach that has been suggested[29,30] is to replace $\mathcal{E}$ with a χ^2-like statistic,[31] such as ϵ^2 (eq 4). In eq 4, the deviation at each data point is normalized

$$\epsilon^2 = \frac{(N_{dip}/\nu)\sum_{i=1}^{N} k^6(\chi_{obs}(k_i) - \chi_{calc}(k_i))^2/\sigma_i^2}{N} \qquad (4)$$

by the variance in that point, σ_i^2, N is the number of data points being fit, and ν represents the number of degrees of freedom ($\nu = N_{idp} - N_{var}$, where N_{var} is the number of variable parameters). N_{idp} is the number of independent data points, given by eq 5, where the energy range of the useful EXAFS is Δk, and ΔR is the region in the Fourier transformed spectrum where the data are physically meaningful.

(29) Lytle, F. W.; Sayers, D. E.; Stern, E. A. *Physica B* **1988**, *158*, 701−722.

(30) Bunker, G.; Hasnain, S.; Sayers, D. In *X-ray Absorption Fine Structure*; Hasnain, S. S., Ed.; Ellis Horwood: New York, 1991; pp 751−770.

(31) Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill Book Company: New York, 1969; pp 187−203.

$$N_{idp} = \frac{2\Delta k \Delta R}{\pi} \quad (5)$$

The inclusion of $1/\nu$ weighting in eq 4 introduces a penalty for adding additional parameters that are not required. Application of eq 4 is difficult in practice as the variance contains both statistical and nonstatistical errors. Although the statistical errors can be determined in a straightforward fashion, the nonstatistical errors, which include systematic errors both in data collection and in the theoretical parameters, are more difficult to determine. One approximation to ϵ^2 is to use $\mathcal{E}'$ (eq 6) or a similar statistic

$$\mathcal{E}' = \mathcal{E}^2/\nu \quad (6)$$

as a measure of fit quality.[32] This amounts to assuming that σ_i^2 in eq 4 is the same for all data points. The improvement in $\mathcal{E}'$ (or in $\mathcal{E}^2$) can then be used in conjunction with the *F*-test[31,33] to determine if the improvement in the fit is statistically significant. In the *F*-test, one calculates the ratio of reduced χ^2 values for the one and two shell fits, $F = [(\mathcal{E}_2^2 - \mathcal{E}_1^2)/2]/[\mathcal{E}_2^2/\nu]$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are for 1 and 2 shell fits. The observed *F* value is compared to the improvement ratio expected simply as a result of the increase in variable parameters. The *F*-test has been used previously in EXAFS analyses,[34] although in the earlier work, the number of measured data points rather than the number of independent data points (N_{idp}) was used calculate *F*, giving erroneously large *F* values. For $N_{idp} \approx 10$, the $\mathcal{E}$ value should decrease by ca. 50% for the two shell fit relative to the single shell fit in order to judge the fit significant at the $P = 0.05$ level ($F > 5.79$).[31,33] The relevant *F* values are given in Table 2. The 2S + 2N fit to the ZnS_2N_2 inorganic model, is, by this criterion, significant at the $P \approx 0.05$ level. However, the 2S + 2N fit to the ZnS_2N_2 peptide model, and the 3S + N fits to either of the ZnS_3N models are *not* significant, despite the fact that these data all unambiguously include both S and N scatterers.

As an alternative to the *F*-test, we have adopted an approach in which a series of two-shell fits were performed for each data set. Each fit uses the same number and type of variable parameters. For each shell, only *R* and σ are allowed to vary, while the scale factor and ΔE_0 are held constant at their respective calibrated values. The total zinc coordination number is held fixed at four and the number of nitrogen scatterers is varied in increments of 0.25. Although a fractional number of sulfur and nitrogen ligands has no physical reality for mononuclear zinc models, it allows for direct comparison of the trends within the different data sets.

The fractional improvement in the *i*th fit is $(\mathcal{E}_{4S} - \mathcal{E}_i)/\mathcal{E}_{4S}$, where $\mathcal{E}_{4S}$ is the mean-square deviation for a one-shell (two variable parameters) fit with four sulfurs and $\mathcal{E}_i$ is the mean-square deviation for the *i*th two-shell (four variable parameters) fit with n_iS + $(4 - n_i)$N. The fractional improvement, plotted as a function of percent sulfur, has a characteristic shape for each different tetracoordinate Zn site, and we have used this dependence to assign Zn site structures.[35−37] In this initial work, we found that one of the important factors for distinguishing between different sites was the extent to which the *n*S + $(4 - n)$N fits improved in comparison to the improvement that is expected simply as a consequence of the increase in variable parameters. The latter can be approximated by determining $\mathcal{E}$ for a fit by using two independent shells of sulfur, $\mathcal{E}_{2S+2S}$. Although the 2S + 2S fits are not physically meaningful, since two independent sulfur shells are not present, they do give a measure of the baseline improvement that results from doubling the number of variable parameters. The excess percent improvement, P_i, is then given by eq 7.

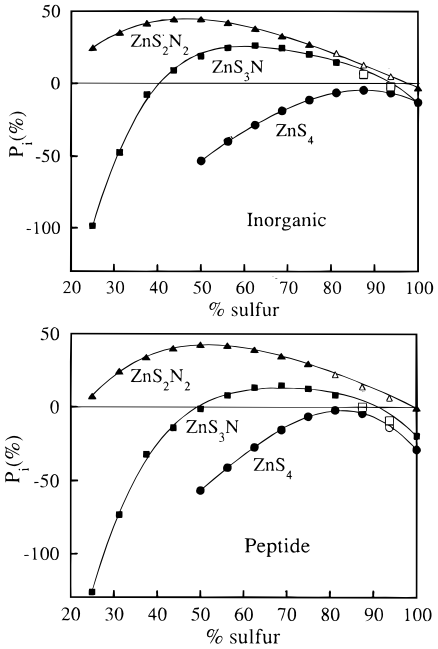

Figure 4. Percent improvement, P_i, as a function of percent sulfur included in the fit: triangles, ZnS_2N_2; squares, ZnS_3N; circles, ZnS_4. The solid curves are third-order polynomial fits to the data: (top) inorganic complexes and (bottom) zinc binding peptides.

$$P_i = \left[\frac{\mathcal{E}_{4S} - \mathcal{E}_i}{\mathcal{E}_{4S}} - \frac{\mathcal{E}_{4S} - \mathcal{E}_{2S+2S}}{\mathcal{E}_{4S}}\right](100\%) = \frac{\mathcal{E}_{2S+2S} - \mathcal{E}_i}{\mathcal{E}_{4S}}(100\%) \quad (7)$$

Figure 4 shows a plot of P_i as a function of the percent sulfur for all of the models studied. The points marked with open symbols have fits for which the minima are not well-defined, giving unreasonable zinc−nitrogen bond lengths if both R_N and σ_N are refined. For these fits, the R_N was held constant at a value extrapolated from other fits in the same series. This phenomenon was only observed for fits having less than 25% nitrogen. These fits are poorly defined as a consequence of the very small contribution that the Zn−N component makes to the overall fit.

Trends in P_i. For all of the models studied, P_i increases as the percent sulfur is decreased below 100%, even for the ZnS_4 models. The P_i vs percent sulfur curves can be characterized by $P_{i,\,max}$, the maximum value found for P_i, and by S_{max}, the percent sulfur giving $P_{i,max}$. As expected, S_{max} coincides approximately with the true percent sulfur in the model, with values of 85 or 90% for ZnS_4, 67 or 75% for ZnS_3N, and 45 or 55% for ZnS_2N_2, where the maxima are determined by fitting a quartic polynomial to the P_i vs percent sulfur plots for the inorganic and peptide models. The $P_{i,max}$ values, which are a measure of how much improvement is seen for S + N fits as

(32) Riggs-Gelasco, P. J.; Mei, R.; Yocum, C. F.; Penner-Hahn, J. E. *J. Am. Chem. Soc.* **1996**, *118*, 2387−99.
(33) Afifi, A. A.; Clark, V. *Computer-Aided Multivariate Analysis*, 2nd ed.; Van Nostrand Reinhold Company: New York, 1990; pp 191−196.
(34) Joyner, R. W.; Martin, K. J.; Meehan, P. *J. Phys. C: Solid State Phys.* **1987**, *20*, 4005−4012.
(35) Garcia, G. A.; Tierney, D. L.; Chong, S.; Clark, K.; Penner-Hahn, J. E. *Biochemistry* **1996**, *35*, 3133−9.
(36) Landro, J. A.; Schmidt, E.; Schimmel, P.; Tierney, D. L.; Penner-Hahn, J. E. *Biochemistry* **1994**, *33*, 14213−20.
(37) Clark, K. Ph.D. Thesis, The University of Michigan, 1994.

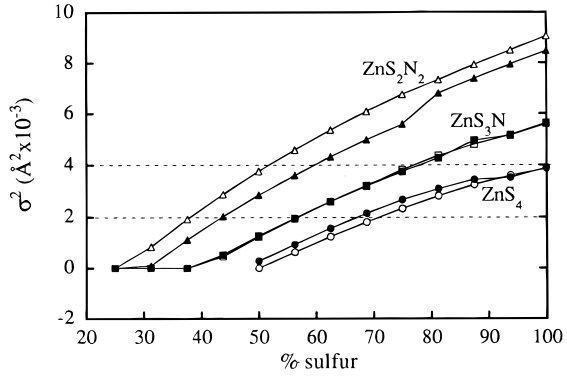

Figure 5. Dependence of σ_S^2 on percent S: triangles, ZnS_2N_2; squares, ZnS_3N; circles, ZnS_4. Open symbols are for the inorganic complexes and filled symbols for the peptides.

compared to S + S fits, increases as the fraction of nitrogen ligands in the model increases, with values slightly less than zero for ZnS_4 models, ~15−25% for ZnS_3N models, and ~40−45% for ZnS_2N_2 models. This trend reflects the fact that the single scatterer (2S + 2S) fits become an increasingly poor description of the data as the number of nitrogen ligands increases. For the ZnS_4 models, the two-shell S + N fits, while better than the one-shell 4S fits, are never better than the two-shell 2S + 2S fits.

The behavior of σ^2. The apparent Debye−Waller factors, σ_S and σ_N, change systematically as the percent S is changed. Figure 5 shows the behavior of σ_S^2 for each of the models studied. In all cases, σ_S^2 increases monotonically as the amount of sulfur used in the fit increases. This is a reflection of the fact that Zn−S scattering dominates the EXAFS; As more sulfurs are used in the fit, a larger σ_S^2 is required to adjust the Zn−S amplitude to match the experimental EXAFS. For a given percent S, the apparent σ_S^2 increases by $(2-4) \times 10^{-3}$ Å^2 on going from S_4 to S_3N to S_2N_2. For all of the models, fits using the true percent sulfur give σ_S^2=$(2-4) \times 10^{-3}$ Å^2. The behavior of σ_N^2 is more complicated, probably as a consequence of the relatively small contribution that Zn−N scattering makes to the overall EXAFS, even for the 2S + 2N fits.

Sensitivity of S_{max} to $S_i(k)$ and ΔE_0. Two of the variables in eq 1, $S_i(k)$ and ΔE_0, are often treated as freely adjustable parameters. In some cases, these are restricted to "chemically reasonable" values, although the range accepted as chemically reasonable can be as large as 0.8 to 1.0 for $S_i(k)$ and ±20 eV for ΔE_0. Plots of S_{max} as a function of the zinc−sulfur scale factor, $S_S(k)$, are unremarkable (see Supporting Information, Figure S1). As $S_S(k)$ increases, S_{max} decreases since fewer Zn−S scatterers are needed to give the same amplitude in the EXAFS fit. For every change of 0.1 in the zinc−sulfur scale factor, there is a 2−4% change in S_{max}. Surprisingly, S_{max}, is also very sensitive to the Zn−S ΔE_0, $\Delta E_{0,S}$, as shown in Figure 6. In general, S_{max} increases by 3−5% for every 1 eV increase in $\Delta E_{0,S}$. It is important to note that although the S_{max} changes as a function of $\Delta E_{0,S}$, the quality of the fits ($\mathcal{G}$) does *not* change significantly as $\Delta E_{0,S}$ is changed, thus making it very difficult to determine the correct ΔE_0 from the goodness of fit. The dependence of S_{max} on the $S_N(k)$ and $\Delta E_{0,N}$ is similar but much smaller (data not shown). S_{max} changes by 1−2% for a 0.1 change in $S_N(k)$ and 0.1−1% per 1 eV change in $\Delta E_{0,N}$.

Sensitivity to the Energy Range. Although the ability to resolve two shells in the Fourier transform depends critically on the energy range of the data (Figure 3), resolution in the FT is *not* necessary to resolve two shells in curve fitting. To study the effect of the energy range on the ability to resolve the S and N shells, several fits were performed with k_{min} varied from 0.5 to 2.0 Å^{-1} and k_{max} varied from 12 to 16 Å^{-1} (typical ranges used in protein data analysis). Neither k_{min} nor k_{max} had a significant effect on either S_{max} or $P_{i,max}$. In particular, the typical k_{max} for protein data, ~12 Å^{-1}, provides adequate sensitivity to distinguish the different Zn sites.

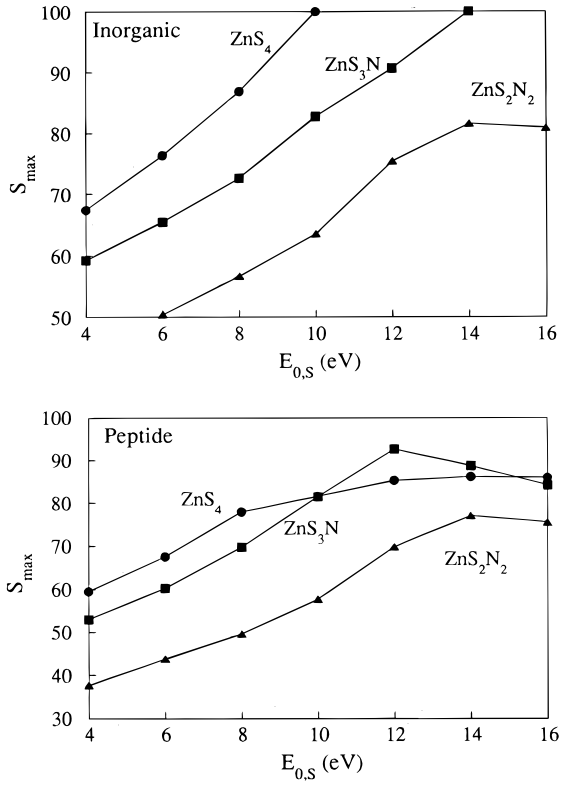

Figure 6. Dependence of S_{max} on Zn−S threshold energy $\Delta E_{0,S}$ for (top) inorganic complexes and (bottom) peptide models: triangles, ZnS_2N_2; squares, ZnS_3N; circles, ZnS_4.

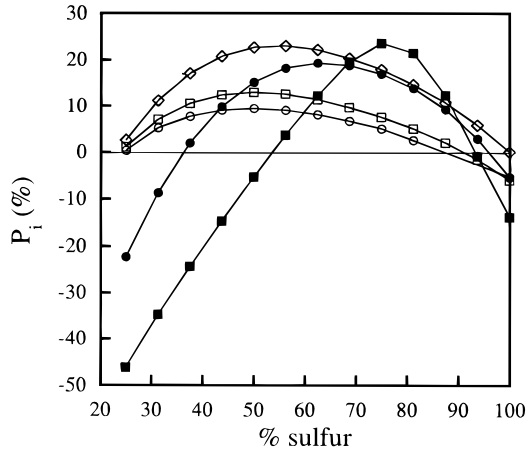

Figure 7. Effect of Fourier filter range on S_{max} for the ZnS_2N_2 peptide: open circles, R = 0.1−3.9 Å; open squares, R = 0.8−3.0 Å; open diamonds, R = 0.8−2.7 Å; filled circles, R = 1.1−2.3 Å; filled squares, R = 1.4−2.2 Å.

Sensitivity to the Filter Range. Fourier filtering is often used to remove high-frequency noise. Figure 7 shows the effect that the Fourier filter range has on P_i for the ZnS_2N_2 peptide model. There is little change in S_{max} as the filter range is decreased to 0.8−2.7 Å. However, when the filter is made smaller (1.1−2.3 or 1.4−2.2 Å) and begins to exclude real data, S_{max} increases, making the data appear consistent with an S_3N site. The relevant filter ranges are shown in Figure S2.

Sensitivity to Noise. Data were measured for both concentrated (~8 mM) and dilute (~1 mM) samples of the peptide models. The noise level for the dilute samples (data not shown) is significantly higher than that for the concentrated samples, although still comparable to that in many published protein EXAFS spectra.[12,14,38−51] The EXAFS data for the dilute peptides gave comparable results to those obtained for the concentrated samples. To evaluate systematically the effect of noise, we compared the results obtained with the average of 8 scans with those obtained with averages of only a subset of scans. Representative spectra are shown in Figure S3. Once again, the low and high noise spectra gave the same $P_{i,max}$ and S_{max} values, and the standard deviation in the $P_{i,max}$ and S_{max} values did not increase as the noise level increased (i.e., as the number of spectra used in the average decreased). These results demonstrate that our method can give accurate results, even for data having a relatively low signal-to-noise ratio.

Discussion

The utility of XAFS for structure elucidation is well established and XAFS has frequently been used to characterize Zn protein sites. However, there have been few attempts to test the validity of the structural models that have been put forward. Where the sensitivity of XAFS *has* been examined,[8−10] it has been concluded that XAFS does not have sufficient sensitivity to distinguish unambiguously between the different possible Zn sites. The objective of the present study was to use a series of structurally characterized Zn models to determine the extent to which different Zn coordination environments can be distinguished. In principle, the XANES shape, the FT features, and the curve-fitting results all contain features necessary to distinguish between ZnS_4, ZnS_3N, and ZnS_2N_2 sites. The utility of each of these is discussed separately below.

XANES Spectra. The inflection point, edge shape, and transition intensities for both sets of model edges change systematically as the ligation changes. In general, the intensity of the transition at 9664 eV increases with increasing thiolate ligation while the transition at ca. 9670 eV, which is broad for the ZnS_2N_2 models (giving two resolved transitions for the inorganic ZnS_2N_2 model), becomes somewhat narrower for tetrathiolate ligation. The broad 9670 eV transition for the ZnS_2N_2 models most likely reflects interference between the first and outer shell scattering from the imidazole ligands. Similar structures have been seen in porphyrin XANES spectra.[52]

The differences between the spectra for the inorganic models and those for the analogous peptides demonstrate that the XANES features depend on more than the scatterer identity. The spectral differences are most likely due to differences in the ligation (e.g. aromatic vs alkanethiolate) and to differences in geometry. While it is clear that zinc XANES features are useful for detecting *changes* in local site structure, the XANES spectra do not contain sufficient information to unambiguously determine the coordination sphere of an unknown zinc site.

The Fourier Transform. It is widely recognized that Fourier transforms are not generally useful for quantitative interpretation of EXAFS spectra. However, Fourier transforms are often used for qualitative interpretation, thus it is instructive that *no* Zn−N shoulder is resolved for $k_{max} \leq 13$ Å^{-1}. The height of the FT modulus scales with the number of thiolate ligands, reflecting the fact that Zn−S scattering dominates the overall EXAFS. Unfortunately, FT height cannot be used to distinguish between ligation types, since disorder can mimic the amplitude decrease that is observed for mixed S + N ligation.

A second way in which the Fourier transform might be used to identify ligation is by identifying outer shell scattering from the imidazole ligands. It is well established that imidazole ligands give rise to pronounced outer shell scattering, with peaks at ~2.5 and 3.5 Å.[53] Quantitation of this outer shell scattering can be used to determine the number of coordinated imidazole ligands. Initial examination of Figure 2 suggests that there is little or no outer shell scattering attributable to the imidazole ligands, even for the ZnS_2N_2 complexes. A more careful comparison between Figure 2 and the FT for $[Zn(N\text{-methylimidazole})_4]^{2+}$ (data not shown) shows close correspondence between the peaks at $R + \alpha \approx 2.8$, 3.3, and 3.6 Å and the outer shell peaks from imidazole ligands. As expected, the outer shell amplitude for the ZnS_2N_2 models in Figure 2 is approximately half as intense as that seen for the tetraimidazole complex. Thus, the correct interpretation of Figure 2 is *not* that the imidazole outer shell scattering is unusually weak, but rather that the first shell scattering is very strong. This is typical of Zn−S scattering.

The FTs for both of the ZnS_4 models show outer shell peaks that have amplitudes comparable to those of the peaks attributable to the imidazole outer shell scattering for the ZnS_2N_2 models. For the ZnS_4 models, the high *R* peaks are most likely due to FT truncation artifacts. Given the results in Figure 2, it seems unlikely that it is possible to uniquely identify imidazole ligation on the basis of the outer shell scattering, even for data having a very good signal/noise ratio. While the presence of imidazole outer-shell scattering needs to be evaluated on a case by case basis, the situation is likely to be even worse for noisy protein data.

Curve Fitting. As discussed previously, one of the difficulties of nonlinear least-squares curve fitting is that an increase in the number of parameters almost always results in an improved fit, regardless of whether the additional parameters are justified. This is illustrated by the fact that the data for *every* zinc model give a lower *R* value for the two-shell 2S + 2S fits than for the one-shell 4S fits. This is true regardless of the actual coordination environment, and despite the fact that

(38) Mangani, S.; Carloni, P.; Viezzoli, M. S.; Coleman, J. E. *Inorg. Chim. Acta* **1992**, *191*, 161−165.
(39) Phillips, J. C.; Bordas, J.; Foote, A. M.; Koch, M. H. J.; Moody, M. F. *Biochemistry* **1982**, *21*, 830−834.
(40) Yachandra, V.; Powers, L.; Spiro, T. G. *J. Am. Chem. Soc.* **1983**, *105*, 6596−6604.
(41) Zhang, K.; Chance, B.; Auld, D. S.; Larsen, K. S.; Vallee, B. L. *Biochemistry* **1992**, *31*, 1159−1168.
(42) Lin, S.-L.; Stern, E. A.; Kalb (Giboa), A. J.; Zhang, Y. *Biochemistry* **1991**, *30*, 2323−2332.
(43) Naqui, A.; Powers, L.; Lundeen, M.; Constantinescu, A.; Chance, B. *J. Biol. Chem.* **1988**, *263*, 12342−12345.
(44) Chance, M. R.; Sagi, I.; Wirt, M. D.; Frisbie, S. M.; Scheuring, E.; Chen, E.; Bess, J. W.; Henderson, L. E.; Arthur, L. O.; South, T. L.; Perez-Alvarado, G.; Summers, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10041−10045.
(45) Freedman, L. P.; Luisi, B. F.; Korszun, Z. R.; Basavappa, R.; Sigler, P. B.; Yamamoto, K. R. *Nature* **1988**, *334*, 543−546.
(46) Garcia-Iniguez, L.; Powers, L.; Chance, B.; Selin, S.; Mannervik, B.; Mildvan, A. S. *Biochemistry* **1984**, *23*, 685−689.
(47) Ross, I.; Binstead, N.; Blackburn, N. J.; Bremner, I.; Diakun, G. P.; Hasnain, S. S.; Knowles, P. F.; Vasak, M.; Garner, C. D. *Springer Ser. Chem. Phys.* **1983**, *27*, 337−341.
(48) Garner, C. D.; Hasnain, S. S.; Bremner, I.; Bordas, J. *J. Inorg. Biochem.* **1982**, *16*, 253−256.
(49) Feiters, M. C.; Jeffery, J. *Biochemistry* **1989**, *28*, 7257−7262.
(50) Diakun, G.; Fairall, L.; Klug, A. *Nature* **1986**, *324*, 698−699.
(51) Navaratnam, S.; Myles, G. M.; Strange, R. W.; Sancar, A. *J. Biol. Chem.* **1989**, *264*, 16067−16071.
(52) Wang, S. Ph.D. Thesis, The Univeristy of Michigan, 1991.
(53) Strange, R. W.; Blackburn, N. J.; Knowles, P. F.; Hasnain, S. S. *J. Am. Chem. Soc.* **1987**, *109*, 7157−62.

none of the models actually contains two resolvable shells of sulfurs. Clearly, an improvement in the mean-square deviation alone is too insensitive to be used to test the validity of adding an additional shell.

The *F*-test, on the other hand, is too restrictive. By this test, the data do not justify inclusion of the nitrogen shell, even for samples which genuinely contain two shells. The reasons for this failure are clear. Use of the *F*-test requires a true χ^2 statistic (e.g., eq 4). However, the more accessible statistic $\mathcal{E}'$ (eq 6) includes systematic errors arising from incomplete background removal and/or inaccurate amplitude and phase functions. Failure to account for these makes $\mathcal{E}'$ larger than it should be, and thus makes relative improvement in $\mathcal{E}'$ smaller than it should be. Therefore, addition of a second shell appears to be unjustified, even for cases where the second shell is present. Rather than introduce ad hoc corrections for these systematic errors, we have developed a protocol in which the number of variable parameters does not change and in which the systematic errors are held constant by treating all data in the same fashion.

In doing this, we held the total coordination number fixed at four. Higher coordination numbers are occasionaly observed for Zn sites and, at least in principle, lower coordination numbers are also possible.[19] It is difficult to use EXAFS amplitudes to determine the absolute coordination number to better than ±1. Fortunately, crystallographic studies have shown that the average zinc−sulfur bond length and the total zinc coordination number are highly correlated. For Zn-thiolate models, the average Zn−S bond length increases from 2.23 Å for ZnS_3 to 2.33 Å for ZnS_4 and 2.47 Å for ZnS_5.[54] Although there is some variation in the bond lengths for mixed ligation sites, the average Zn−S distance remains close to these values. Since EXAFS can be used to determine the average bond lengths with very high accuracy (±0.02 Å), the average Zn−S and Zn−N distances, in combination with bond-valence-sum calculations,[55,56] can be used to determine the total zinc coordination number with a high degree of confidence.

The results shown in Figure 4 illustrate the sensitivity of our approach for characterizing ZnS_xN_y ligation. There are characteristic, reproducible differences in the shapes of the P_i vs percent S curves for each of the different ligation types. As the fraction of N ligands increases, $P_{i,max}$ increases, reflecting the increased need for both Zn−S and Zn−N scattering to model the data. Similarly, S_{max} decreases as the fraction of S ligands decreases, such that S_{max} is close to the true percent S. A striking feature of Figure 4 is the fact that even the ZnS_4 models show moderate improvements in fit quality on inclusion of a small percentage of nitrogen ligation, giving $S_{max} \approx 85\%$. This increase could easily result in mistakenly identifying a ZnS_4 site as having some fraction of low-Z ligation (e.g., ZnS_3N). However, by controlling the number of degrees of freedom (i.e., by referencing P_i to the improvement that is seen for 2S + 2S fits), it is straightforward to demonstrate that no significant improvement is seen for the S + N fits to the ZnS_4 data. Only for the true mixed ligand sites is $P_{i,\,max}$ larger than zero.

In addition to P_i, σ^2 is also useful for distinguishing between Zn environments, as illustrated in Figure 5. The dependence of σ_S^2 on percent S simply reflects the fact that Zn−S scattering can be made more "nitrogen-like" by increasing σ_S^2, while Zn−N scattering can be made more "sulfur-like" by decreasing σ_N^2. On the basis of the models studied, reasonable values for σ_S^2 are ca. $(2.0-4.0) \times 10^{-3}$ Å^2. This is probably a reasonable lower limit for σ_S^2, but cannot be used as an upper limit, since disorder in the Zn−S shell could cause genuine increases in σ_S^2 beyond the values shown in Figure 5.

Perhaps the most striking result was the finding that small variations in E_0 can result in quite large variations in the apparent percent S. In retrospect, this observation makes sense, since small differences in the EXAFS phase (which is defined by E_0) can have significant effects on the interference between the Zn−S and the Zn−N shells. In many EXAFS analyses, E_0 is treated as a freely variable parameter. Although E_0 is often constrained to fall within "chemically reasonable" limits, these have been defined to be as large as 10 eV[57] or even larger.[58] In some cases, different E_0 values are refined for each different shell.[44,59,60] The results in Figure 6 demonstrate that even small errors in E_0 can result in large errors in the apparent structure and could easily result in misidentification of the ligation. The safest procedure is to calibrate E_0 with EXAFS data for structurally similar model compounds (i.e., the value of 9 eV used in the present study). If E_0 is allowed to vary, the limits for variation should be small and E_0 for each shell should be the same. Fits in which the variation of E_0 relative to the calibrated value, $|\Delta E_0|$, is greater than approximately 2 eV or in which different shells refine to different E_0 values should be regarded with suspicion.

A final point concerns the importance of Fourier filter range. The use of narrow filters can profoundly affect the apparent structure, as determined by EXAFS. Conceptually the problem is straightforward: A filter range that is too narrow can eliminate frequencies corresponding to genuine structural contributions. In practice, however, it is important to recognize that relevant FT contributions include shoulders, and in some cases side lobes on the main peak(s).

Summary

The relative proportions of sulfur and nitrogen scatterers around a zinc site can be determined reliably with X-ray absorption spectroscopy. While the normalized XANES spectra and the Fourier transform afford some insight into the zinc ligand sphere, neither can be used to determine the coordination site unambiguously. Quantitative curve fitting analysis provides the further evidence needed to distinguish between different zinc ligation possibilities. However, improvement in fit quality alone is not sufficient evidence for the presence of multiple shells of scatterers. Only by correcting for the number of variable parameters is it possible to assign the ligation reliably.

The phase and amplitude functions used in quantitative curve fitting should be calibrated against model compounds resembling the unknown as much as possible, and the resulting calibrated parameters (scale factor and E_0) should be held fixed for subsequent refinements. If these are allowed to vary, misleading results are likely. Although we have only demonstrated this for tetrahedral Zn sites, this conclusion is likely to be applicable to most EXAFS analyses. When considered together, the values of $P_{i,max}$, S_{max}, and σ^2 allow unambiguous discrimination between the different possible zinc environments. The calibration procedure developed here can be readily extended to other

(54) Average bond lengths are based on a survey of Zn thiolate structures deposited in the Cambridge Crystallographic Data Bank.

(55) Brown, I. D.; Altermatt, D. *Acta Crystallogr.* **1985**, *B41*, 244−247.
(56) Thorp, H. H. *Inorg. Chem.* **1992**, *31*, 1585−8.
(57) Chance, M.; Powers, L.; Kumar, C.; Chance, B. *Biochemistry* **1986**, *25*, 1259−1265.
(58) Liang, W.; Latimer, M. J.; Dau, H.; Roelofs, T. A.; Yachandra, V. K.; Sauer, K.; Klein, M. P. *Biochemistry* **1994**, *33*, 4923−4932.
(59) Dau, H.; Andrews, J. C.; Roelofs, T. A.; Latimer, M. J.; Liang, W.; Yachandra, V. K.; Sauer, K.; Klein, M. P. *Biochemistry* **1995**, *34*, 5274−87.
(60) Sinclair, R.; Hallam, S.; Chen, M.; Chance, B.; Powers, L. *Biochemistry* **1996**, *35*, 15120−15128.

coordination numbers, to other mixed ligand environments such as P/S + As/Se, and to other absorbing atoms.

Acknowledgment. This work was supported in part by grants from the NIH (GM-38047 to J.P.H. and GM31849 to S.A.K.). SSRL and NSLS are supported by the US Department of Energy with additional support from the NIH Research Resource program. We thank Katrina Peariso for helpful discussions.

Supporting Information Available: Figures showing the effect of sulfur scale factor on S_{max} (Figure S1), the R ranges for the filters used to produce Figure 7 (Figure S2), and the "low" and "high" noise data used to test the sensitivity of the analysis protocol to noise (Figure S3) (3 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA980580O